

## Methylation of *HpaII* and *HhaI* Sites Near the Polymorphic CAG Repeat in the Human Androgen-Receptor Gene Correlates with X Chromosome Inactivation

R. Cutler Allen,\* Huda Y. Zoghbi,\*† Annemarie B. Moseley,\* Howard M. Rosenblatt,† and John W. Belmont\*†‡

\*Institute for Molecular Genetics, †Department of Pediatrics, and ‡Howard Hughes Medical Institute, Baylor College of Medicine, Houston

### Summary

The human androgen-receptor gene (HUMARA; GenBank) contains a highly polymorphic trinucleotide repeat in the first exon. We have found that the methylation of *HpaII* and *HhaI* sites less than 100 bp away from this polymorphic short tandem repeat (STR) correlates with X inactivation. The close proximity of the restriction-enzyme sites to the STR allows the development of a PCR assay that distinguishes between the maternal and paternal alleles and identifies their methylation status. The accuracy of this assay was tested on (a) DNA from hamster/human hybrid cell lines containing either an active or inactive human X chromosome; (b) DNA from normal males and females; and (c) DNA from females showing nonrandom patterns of X inactivation. Data obtained using this assay correlated substantially with those obtained using the PGK, HPRT, and M27 $\beta$  probes, which detect X inactivation patterns by Southern blot analysis. In order to demonstrate one application of this assay, we examined X inactivation patterns in the B lymphocytes of potential and obligate carriers of X-linked agammaglobulinemia.

### Introduction

Dosage compensation in humans is achieved through random inactivation of one of the two X chromosomes in the cells of normal females (Gartler and Riggs 1983; Lyon 1989). This inactivation occurs early in development and results in individuals who are essentially cellular mosaics with either the maternal or paternal X chromosome inactivated (Lyon 1972). Reduction of gene expression is believed to result directly from the inactivation process, and this affects all regions of the X chromosome, with only a few exceptions (Gartler and Riggs 1983; Grant and Chapman 1988; Brown and Willard 1990). Methylation of deoxycytosine residues is hypothesized to be a factor involved

in the modulation of gene expression, and the 5' regions of genes, especially the promoters and GC-rich CpG islands, appear to be the targeted areas for methylation (Wolf et al. 1984; Yen et al. 1984; Keith et al. 1986; Hansen et al. 1988; Toniolo et al. 1988). As of yet, the sequence of events that leads to methylation is unknown; however, recent discovery of the XIST gene in the X inactivation center promises to bring greater understanding of the fundamental mechanisms of X inactivation (Brown et al. 1991a, 1991b).

Determination of maternal and paternal X chromosome activation status is useful in the diagnostic analysis of nonrandom X inactivation patterns. This has been of interest in studies of clonality of neoplastic cells in females (Vogelstein et al. 1985; Fearon et al. 1987a) and for evaluating skewed patterns of inactivation in particular tissues from female carriers of a number of X-linked recessive diseases (Puck et al. 1987; Goodship et al. 1988; Greer et al. 1989), females who are symptomatic carriers of an X-linked recessive disease (Gomez et al. 1977), and females who are asymptomatic carriers of an X-linked dominant disease

Received March 20, 1992; final revision received August 5, 1992.

Address for correspondence and reprints: Dr. John W. Belmont, Baylor College of Medicine, One Baylor Plaza, T828, Houston, TX 77030.

© 1992 by The American Society of Human Genetics. All rights reserved. 0002-9297/92/5106-0006\$02.00

(Migeon et al. 1981). One method of analyzing X inactivation takes advantage of the methylation-sensitive restriction enzymes *HpaII* and *HhaI*. The most common probes used in this type of assay are the 5' regions of the HPRT and PGK genes (Wolf et al. 1984; Yen et al. 1984; Vogelstein et al. 1987) and, more recently, the anonymous marker M27 $\beta$  (Boyd and Fraser 1990; Brown et al. 1990). With each of these probes, an RFLP (HPRT and PGK) or VNTR (M27 $\beta$ ) is adjacent to a *HpaII* or *HhaI* enzyme site, and methylation of the site correlates with X inactivation. This enables the methylation status of both the maternal and paternal X chromosomes to be determined in an assay using Southern analysis or PCR (Singer-Sam et al. 1990). Another method used in detecting nonrandom X inactivation patterns involves the production of hamster/human hybrids (Puck et al. 1987). This method is limited to cells that can be fused with HPRT-negative hamster cells. The hybrids are selected for HPRT production and, therefore, the active human X chromosome. The hybrids readily lose the human chromosomes and will only consistently retain a human chromosome under selection for a gene on that chromosome. When DNA from individual HPRT-positive hybrid clones is analyzed, with X-linked probes, the ratio of hybrids retaining the maternal or paternal X chromosome as the active X chromosome indicates the pattern of X inactivation.

In this report we show that the methylation of *HpaII* and *HhaI* sites in the first exon of the human androgen-receptor locus correlates with X inactivation. The close proximity of these enzyme cleavage sites to a highly polymorphic (20 alleles, 90% heterozygosity) short tandem repeat (STR) allows the use of a PCR assay to identify the methylation patterns of the maternally and paternally derived X chromosomes. Samples tested with the assay include DNA from normal males and females, DNA from hamster/human hybrids that contain either an active or inactive human X chromosome, and DNA from females who show skewed X inactivation patterns. The data obtained using the described PCR assay correlated substantially (22 of 24 samples) with those obtained from the established assays that use Southern analysis. To illustrate one of the many potential applications of this assay, we examined the X inactivation patterns in the B lymphocytes of carriers of X-linked agammaglobulinemia (XLA). In carriers for this disorder, there is skewing of X inactivation in the B lymphocytes (Conley et al. 1986; Fearon et al. 1987b; Conley and Puck 1988). A family with a history of XLA was studied, and pre-

dictions of carrier status that were made on the basis of (a) the independent results of the androgen-receptor assay as a carrier test and (b) linkage analysis correlated completely.

## Subjects and Methods

### Subjects

Tissue samples were obtained from informed and consenting individuals. A total of 6 normal males, 10 normal females, 24 females with known X inactivation patterns as assessed previously by established procedures, and 13 individuals from a family with a history of XLA were studied.

### PCR

DNA from hamster/human hybrid cell lines containing either an active or inactive human X chromosome were obtained from Dr. A. C. Chinault (Institute for Molecular Genetics, Baylor College of Medicine). Human DNA was isolated from peripheral blood by standard procedures (Maniatis et al. 1982). For each DNA sample, three reactions were prepared: in one, 2  $\mu$ g of DNA was digested with 20 U *HpaII* (Pharmacia); in another, 2  $\mu$ g of DNA was digested with 20 U *HhaI* (Pharmacia); in the last, DNA was incubated with the enzyme digestion buffer with no enzyme. All reactions were 20  $\mu$ l in total volume, and all incubations were for 12 h at 37°C. After digestion, the reactions were terminated by incubating the mixture at 95°C for 10 min. From this reaction, 2  $\mu$ l was added to 30  $\mu$ l total volume PCR reaction containing both oligonucleotide primers at a concentration of 1  $\mu$ M, 250  $\mu$ M dNTPs (Pharmacia), 0.05 U *Taq* polymerase/ $\mu$ l (Perkin-Elmer), 0.134 nCi  $\alpha$ -<sup>32</sup>P dCTP/ $\mu$ l (>3,000 Ci/mmol; ICN), 2.5 mM MgCl<sub>2</sub>, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), and 0.01% (w/v) gelatin. Oligonucleotides were synthesized on an ABI 380B oligonucleotide synthesizer by the nuclear acid core facility of the Institute for Molecular Genetics at Baylor College of Medicine. The sequences of the primers were obtained from sequences reported elsewhere (Tilley et al. 1989): primer 1, 5'-GCTGTGAA-GGTTGCTGTTCTCAT-3'; and primer 2, 5'-TCC-AGAATCTGTTCCAGAGCGTGC-3'. Samples were amplified using a Perkin-Elmer thermocycler for 28 cycles (each comprising 45 s at 95°C, 30 s at 60°C, and 30 s at 72°C), with an initial denaturation at 95°C for 5 min. Two microliters of the PCR product was mixed with 5  $\mu$ l of sequencing gel-loading buffer (98% deionized formamide, 10 mM EDTA [pH 8.0],

0.025% xylene cyanol FF, and 0.025% bromophenol blue). One and a half microliters of this mixture was loaded on a denaturing 6% 39:1 acrylamide/bis-acrylamide gel (8 M urea and 1 × TBE), and electrophoresis was performed at 80 W for 3 h. The gel was dried and exposed to X-ray film (X-OMAT AR; Kodak) at  $-80^{\circ}\text{C}$  for 12 h with an intensifying screen. Autoradiographs were then scored blindly.

#### Southern Hybridization

Human DNA was isolated by standard procedures (Maniatis et al. 1982) from peripheral blood. HPRT, PGK, and M27 $\beta$  heterozygosity were identified according to methods reported elsewhere (Wolf et al. 1984; Yen et al. 1984; Fraser et al. 1987; Vogelstein et al. 1987). Enzymes used were obtained from Boehringer-Mannheim. The HPRT and PGK probes were provided by Bert Vogelstein (Johns Hopkins University); the M27 $\beta$  probe was provided by Ian Craig (University of Oxford). Probes were labeled with  $\alpha$ - $^{32}\text{P}$  dCTP ( $>3,000\text{Ci}/\text{mmol}$ ; ICN) to a specific activity of  $>5 \times 10^8 \text{ cpm}/\mu\text{g}$  by using random hexamer priming and extension with the Klenow fragment of DNA polymerase I (Pharmacia) (Feinberg and Vogelstein 1983). Southern hybridization was performed using Nytran (Schleicher and Schuell) membranes. In order to determine X inactivation patterns for the HPRT, PGK, and M27 $\beta$  probes, procedures described elsewhere (Wolf et al. 1984; Yen et al. 1984; Vogelstein et al. 1987; Boyd and Fraser 1990; Brown et al. 1990) were used. All autoradiographs were scored blindly.

#### Carrier Analysis for XLA

B lymphocytes were isolated from the peripheral blood of individuals by using CD19 immunomagnetic beads (Dynal, Inc.) according to a method described elsewhere (Moseley and Huston 1991). The estimated purity by flow cytometry was 95%–98%. Pelleted B lymphocytes ( $2.6 \times 10^5$ ) were resuspended in 75  $\mu\text{l}$  of lysis buffer (0.1 mg proteinase K/ml, 3.4  $\mu\text{M}$  SDS, 40 mM DTT, 2.5 mM  $\text{MgCl}_2$ , 50 mM KCl, 10 mM Tris-HCl [pH 8.3], and 0.01% gelatin) and were incubated at  $37^{\circ}\text{C}$  for 4 h. The reaction was then incubated at  $85^{\circ}\text{C}$  for 10 min. For each B lymphocyte sample, two incubations were prepared: in one, 5  $\mu\text{l}$  of the lysis reaction was digested with 10 U *Hpa*II (Pharmacia); and, in the other, 5  $\mu\text{l}$  of the lysis reaction was incubated with the enzyme digestion buffer and no enzyme. The reactions were 10  $\mu\text{l}$  total volume and incubated at  $37^{\circ}\text{C}$  for 12 h. The reactions were terminated at  $95^{\circ}\text{C}$  for 10 min. From the reactions,

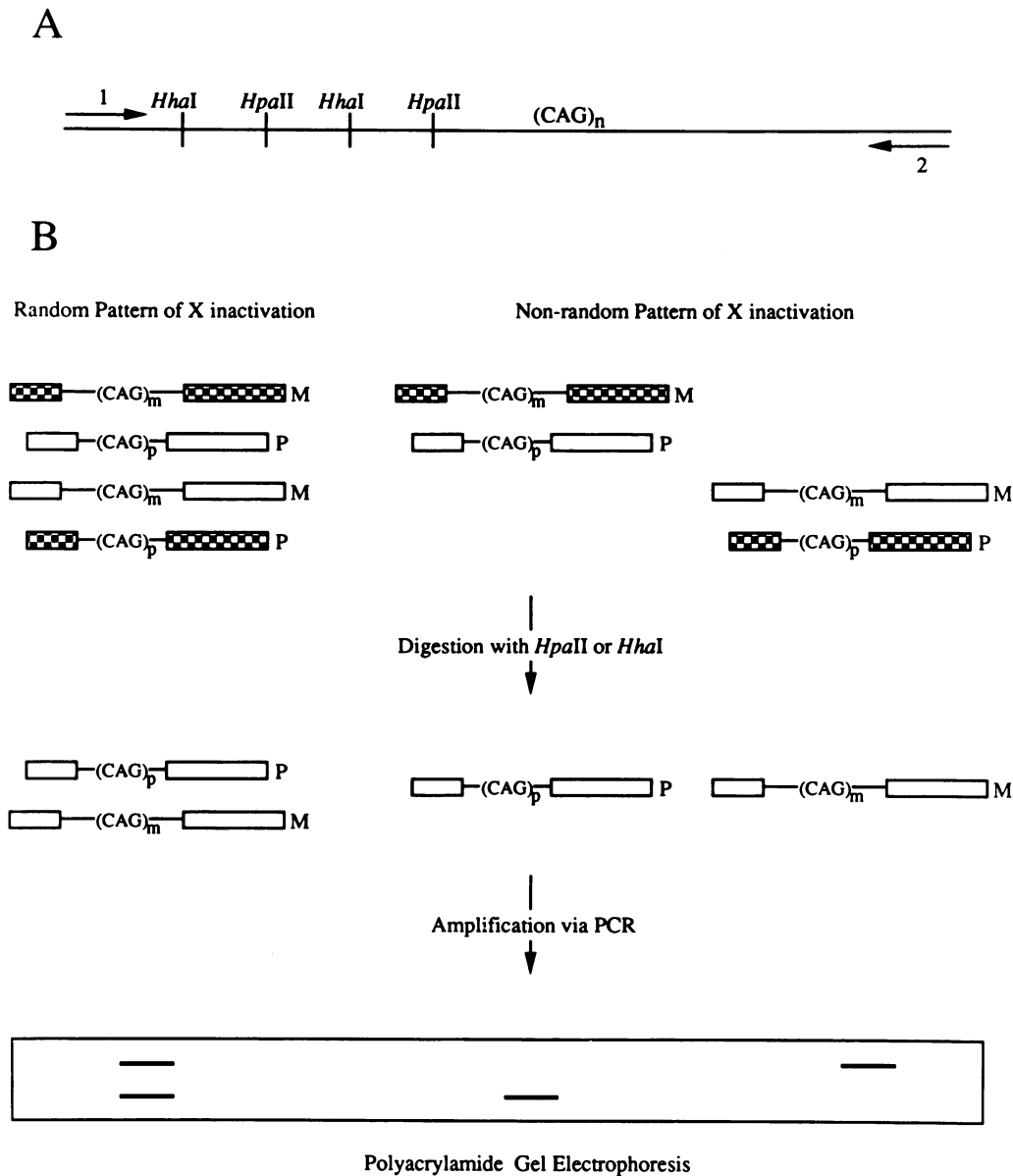
5  $\mu\text{l}$  was added to a 20- $\mu\text{l}$  total volume PCR reaction with concentrations and conditions as noted above. The only difference was the absence of  $\text{MgCl}_2$  from the PCR reaction buffer; this is provided by the enzyme digestion buffer in the 5  $\mu\text{l}$  that is amplified from the digestion reaction. To control for skewing arising by stochastic mechanisms, CD19 $^{-}$  cells were analyzed. These CD19 $^{-}$  cells consist predominantly of T lymphocytes.

#### Results

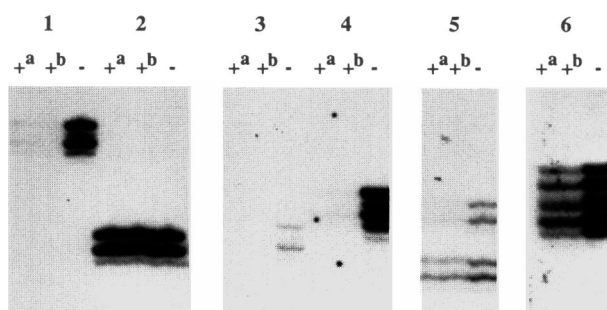
A 280-bp PCR amplification unit including the flanking *Hpa*II and *Hha*I sites and the trinucleotide repeat element (nucleotides 229–508 [HUMARA; GenBank]; accession number M21748) was developed for the human androgen-receptor locus. When the template DNA is digested initially with either *Hpa*II or *Hha*I, amplification will only occur if the restriction sites are methylated; if any of the sites are unmethylated, digestion will occur between the flanking oligonucleotides, and amplification will not be possible. If methylation of these sites correlates with X inactivation, a product will be obtained only when the X chromosome is inactive. The close proximity of the repeat to the methylation-sensitive restriction-enzyme sites enables the maternal and paternal X chromosomes in a female to be distinguished from each other. In a female in which nonrandom X inactivation is present, the more commonly inactivated allele will be preferentially amplified, and this will be detected by a stronger band appearing for that allele (fig. 1).

In order to verify that the methylation status of the restriction-enzyme sites correlated with X inactivation, the assay was performed on DNA from hamster/human hybrid cell lines that contain either a single active or inactive human X chromosome. The DNA was digested with one of the methylation-sensitive enzymes, followed by PCR across the region. One would predict that the hybrid with the inactive X chromosome would generate a product while the hybrid with the active X chromosome would be digested by the restriction enzymes and fail to amplify. As seen in figure 2 (samples 1 and 2), a PCR product was not obtained from the hybrid that contained an active X chromosome, while a product was obtained from the hybrid that contained an inactive X chromosome.

To test the assay on heterozygous and hemizygous cells, DNA isolated from blood samples of 6 normal males and 10 normal females was analyzed using the assay. The DNA from males did not amplify after



**Figure 1** A, Diagram of the region amplified in the first exon of the human androgen receptor. Two *HhaI* and two *HpaII* sites are within 100 bp 5' to the polymorphic CAG repeat. Methylation of these enzyme sites correlates with X inactivation. Oligonucleotide primers 1 and 2 are designed to flank both the methylation-sensitive restriction-enzyme sites and the CAG repeat simultaneously. Sequences for these primers are listed in Subjects and Methods. B, Analysis of X chromosome inactivation patterns by PCR in the first exon of the human androgen receptor locus. The flow diagram illustrates expected results from DNA isolated from cell populations showing either random (left) or nonrandom (right) X inactivation patterns. M and P = maternal and paternal X chromosomes, respectively; (CAG)<sub>m</sub> and (CAG)<sub>p</sub> = allele associated with the polymorphic CAG repeat on the maternal and paternal X chromosomes, respectively. After incubation with the methylation-sensitive enzymes *HpaII* or *HhaI*, the sites on the active X chromosome (checked) will be cleaved, since they are unmethylated; the sites on the inactive X will not be cleaved, since they are methylated. Amplification by PCR between these primers will only yield a product from the uncleaved inactive X chromosome. The X inactivation patterns are therefore assessed in a female who is informative at the CAG repeat. The maternal and paternal alleles are resolved using PAGE (polyacrylamide gel electrophoresis). The HUMARA alleles are shown as single bands, for graphic clarity. In practice, each allele is represented by two major and two or more minor bands (Edwards et al. 1992).

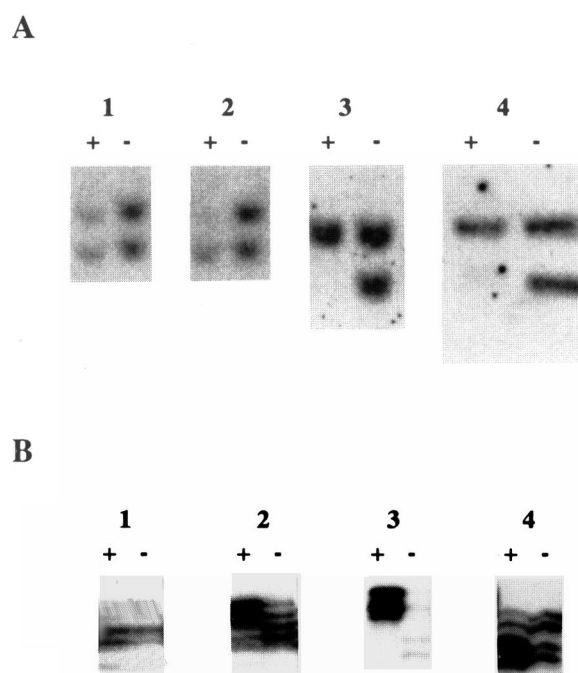


**Figure 2** Analysis of methylation of *HpaII* and *HhaI* sites at the human androgen-receptor locus. Hamster/human hybrid contains an active human X chromosome (gels 1), hamster/human hybrid containing an inactive human X chromosome (gels 2), normal males (gels 3 and gels 4), and normal females (gels 5 and gels 6). Lanes +, DNA amplification with *HpaII* predigestion. Lanes +, DNA amplification with *HhaI* predigestion; Lanes -, DNA amplification without predigestion.

enzyme digestion (fig. 2, samples 3 and 4). This result supports the concept that the active X chromosome is unmethylated, as suggested by the hamster/human hybrid analysis. DNA from normal females who were heterozygous at the CAG repeat amplified at both alleles after enzyme digestion, as indicated by four bands. Each allele is represented by two major bands and several of lesser intensity. Because internal  $^{32}\text{P}$ -dCTP-incorporation labeling is used, both strands of DNA are detected (i.e., each single-stranded DNA migrates slightly differently in the gel, because of complementary base composition). The lighter-intensity bands are presumed to arise in the PCR as multiples of the repeat unit, presumably caused by slippage of the polymerase during PCR. In one sample (fig. 2, sample 6), the intensity of the two alleles is approximately equal, indicating random X inactivation. In the other female DNA sample (fig. 2, sample 5), the intensity of one of the alleles is greater than that of the other allele, indicating skewed X inactivation in an otherwise normal female.

DNA samples were also characterized from 24 females who had been previously studied by Southern analysis with at least one or more of the following probes: PGK, HPRT, and M27 $\beta$ . By Southern analysis, these samples were known to have random or skewed patterns of X inactivation and were chosen accordingly, to provide a substantial base for comparison of the PCR and Southern techniques. With each of the probes used, an RFLP (HPRT and PGK) or VNTR (M27 $\beta$ ) is closely associated with either a *HpaII* or *HhaI* restriction site, and the methylation of

these sites has previously been shown to correlate with X inactivation (Wolf et al. 1984; Yen et al. 1984; Vogelstein et al. 1987; Boyd and Fraser 1990; Brown et al. 1990). These assays require two separate restriction digests. In the first digest, the DNA is incubated with the enzyme(s) that detects the RFLP or VNTR; in the second reaction, the DNA is digested with both the enzyme(s) that detects the RFLP or VNTR and the methylation-sensitive restriction enzyme. The first digestion detects the alleles present at the locus, while the second detects the methylation status of each of the alleles in the DNA sample. With the HPRT and PGK probes, the methylated allele is contained on the inactive X chromosome (Wolf et al. 1984; Yen et al. 1984; Vogelstein et al. 1987); however, with the M27 $\beta$  probe, the methylated allele is contained on the active X chromosome (Body and Fraser 1990). As shown in figure 3A, DNA samples from females



**Figure 3** Analysis of X inactivation patterns by using the reported PCR assay and the PGK probe. X inactivation patterns were studied on DNA samples from the peripheral blood of females. Four examples are shown here: gels 1, gels 2, gels 3, and gels 4 refer to samples 18, 19, 4, and 3, respectively (see table 1). A, Analysis of X inactivation by using the PGK probe. Lanes +, Southern analysis on DNA digested with *BstXI*, *PstI*, and *HpaII*. Lanes -, Southern analysis on DNA digested with *BstXI* and *PstI*. B, Analysis of X inactivation on the same samples shown in A, using the reported PCR assay associated with the human androgen-receptor locus. Lanes +, PCR after *HpaII* predigestion. Lanes -, PCR without predigestion.

exhibiting both random and nonrandom X inactivation patterns were analyzed using these probes. In a DNA sample with random X inactivation (fig. 3A, sample 1), the double-digested DNA shows two bands of equal intensity from each allele, corresponding to the allele cut and uncut with the methylation-sensitive enzyme. In a sample showing nonrandom X inactivation (fig. 3A, samples 2–4), one of the alleles is predominantly digested while the other allele is chiefly undigested. Table 1 lists each of the DNA samples studied and whether the sample exhibited a nonrandom or random pattern of X inactivation as detected by PGK, HPRT, or M27 $\beta$ . One of the samples (table 1, sample 6) was examined by hybrid analysis exclusively.

The samples discussed above were also studied using the androgen-receptor assay. Methylation patterns were scored by analyzing samples with and without *HpaII* or *HhaI* digestion (fig. 3B), if the individual

was heterozygous at the locus. The alleles in the undigested sample were always of approximately equal intensity; methylation patterns were determined by comparing the intensity of the two alleles in samples that had been digested with the restriction endonucleases. In a situation of random X inactivation (fig. 3B, sample 1), the alleles were still of equal intensity after digestion. If nonrandom X inactivation was present (fig. 3B, samples 2–4), one of the alleles was more intense than the other, after digestion. The results obtained in this assay correlated substantially with those from PGK, HPRT, and M27 $\beta$  probes (fig. 3 and table 1). Of the 24 samples selected for analysis that were informative at the androgen-receptor locus, 9 showed nonrandom X inactivation with both the androgen-receptor assay and one or more of the Southern blot assays, and 13 showed random X inactivation with both the androgen-receptor assay and one or more of the Southern blot assays. Two samples, however, gave conflicting results on both Southern analysis assays and the androgen-receptor assay. With sample 22 (table 1), results obtained with the PGK probe show essentially random X inactivation, while the androgen-receptor assay gives an ~75:25 pattern of nonrandom X inactivation. With sample 24 (table 1), results obtained using the M27 $\beta$  probe show an ~80:20 pattern of nonrandom X inactivation, while the androgen-receptor assay gives an essentially random X inactivation pattern. Therefore, of the 24 samples studied, 22 gave identical results by both the androgen-receptor assay and the Southern analysis assays.

A family with a history of XLA was analyzed using the androgen-receptor locus assay (fig. 4). Carriers for XLA show essentially 100% skewing of X inactivation in the B lymphocytes (Conley et al. 1986; Fearon et al. 1987b; Conley and Puck 1988), because of a developmental defect in precursor cells expressing the mutant allele. By isolating the DNA from the CD19<sup>+</sup> cells (93%–96% B lymphocytes) of potential and obligate carriers in the family, we were able to detect X inactivation patterns in the B lymphocyte population. The DNA from the CD19<sup>−</sup> cells (which are predominantly T lymphocytes) of these individuals was also analyzed, to control for skewing arising by stochastic mechanisms; carriers for XLA have normal T cell development and do not exhibit skewing of X inactivation in their T cells (Conley et al. 1986; Fearon et al. 1987b; Conley and Puck 1988). The three obligate carriers (II-2, III-2, and III-4; fig. 4B) all show skewing in their B cells and no skewing in their T cells. Two potential carriers (IV-2 and IV-5) also show skewing in their B

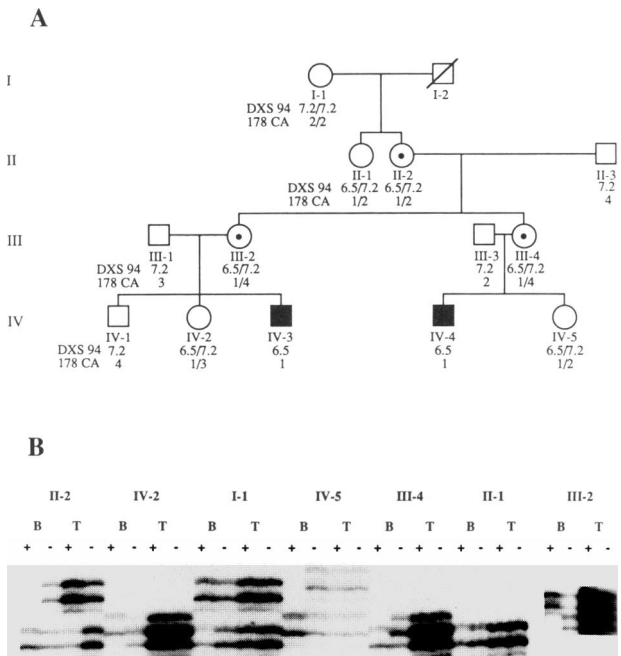
**Table 1**

**X Inactivation Patterns Observed by Using Southern Analysis and the Androgen Receptor Assay**

DNA	PGK	HPRT	M27 $\beta$	HUMARA
1 .....	NR	...	...	NR
2 .....	R	R	...	R
3 .....	NR	...	...	NR
4 .....	NR	...	NR	NR
5 .....	NR	...	...	NR
6 <sup>a</sup> .....	...	...	...	NR
7 .....	NR	...	...	NR
8 .....	R	R	...	R
9 .....	R	...	...	R
10 .....	...	...	R	R
11 .....	R	...	...	R
12 .....	R	...	...	R
13 .....	...	...	R	R
14 .....	...	R	...	R
15 .....	...	...	NR	NR
16 .....	R	...	...	R
17 .....	R	...	...	R
18 .....	R	...	...	R
19 .....	NR	...	...	NR
20 .....	R	...	...	R
21 .....	R	...	...	R
22 .....	R	...	...	NR
23 .....	...	...	NR	NR
24 .....	...	...	NR	R

NOTE.—NR = nonrandom X inactivation; and R = random X inactivation. An ellipsis ( . . . ) indicates that the sample was either uninformative or not analyzed with that probe.

<sup>a</sup> Examined by hybrid analysis only.



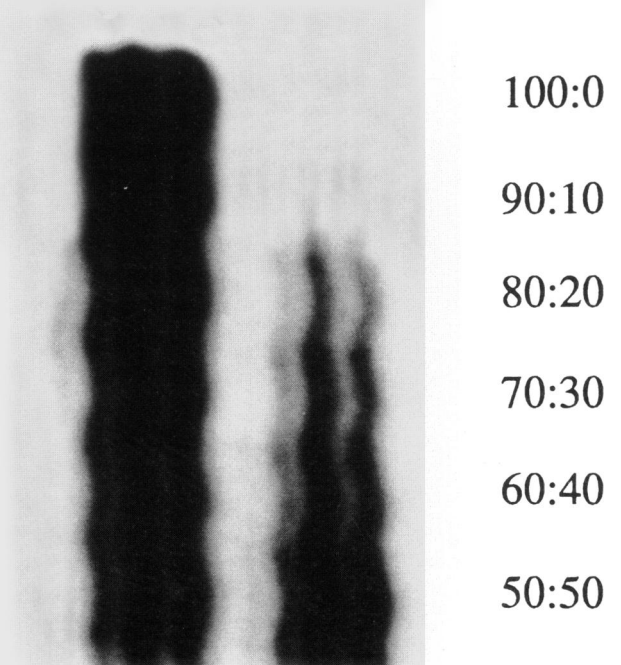
**Figure 4** A, Family with a history of X-linked agammaglobulinemia (XLA). Affected individuals are denoted by blackened squares; and obligate carriers are denoted by circles with a dot. Allele designations for DXS94 and 178 CA (see text) are shown below the number associated with each individual. B, Analysis of X inactivation patterns in the CD19<sup>+</sup> and CD19<sup>-</sup> cells of potential and obligate carriers of XLA. Four lanes are associated with each individual: the first two lanes B designate DNA isolated from the CD19<sup>+</sup> cells (predominantly B lymphocytes) of the individual, and the two lanes T designate DNA from the CD19<sup>-</sup> cells (predominantly T lymphocytes) of the individual. Lanes +, *HpaII* predigestion. Lanes -, No predigestion.

cells and no skewing in their T cells. One potential carrier (I-1) shows no skewing in the B cells or T cells, and another potential carrier (II-1) is uninformative for the CAG repeat. By the carrier test alone, individuals II-2, III-2, III-4, IV-2, and IV-5 are carriers for XLA; individual I-1 is not a carrier for XLA; and no prediction can be made for individual II-1.

The family was also studied by linkage analysis, to assess carrier status. XLA has been mapped to Xq21.3-22 by linkage analysis (Kwan et al. 1986). Two probes, DXS94 and DXS178, have been found to be tightly linked to the locus (Malcolm et al. 1987; Kwan et al. 1990). We have also identified a polymorphic CA repeat from a yeast artificial chromosome (YAC) that contains DXS178 (Allen and Belmont 1992), and this repeat has been found to be useful in the analysis of families with XLA (M. E. Conley, personal communication). When the family was ana-

lyzed with these three probes, the *TaqI* polymorphism of DXS178 was found to be entirely uninformative for the family, while DXS94 and DXS178 CA repeat were both informative in most of the individuals (fig. 4A). In this pedigree, XLA is associated with the 6.5 allele from DXS94 and with the 1 allele from the CA repeat. When the linkage information is used to predict carrier status, individuals II-2, III-2, III-4, IV-2, and IV-5 are carriers, and individual I-1 is not a carrier. The mutation appears to have originated from individual I-2; therefore, individual II-1 is not a carrier, unless there is gonadal mosaicism in individual I-2. When the data from the carrier analysis, which assessed X inactivation patterns, are compared with data from linkage analysis, the results correlate, with no inconsistencies.

In order to demonstrate the quantitative potential of the described assay, a reconstruction experiment was devised by mixing purified CD19<sup>+</sup> cells from two males who had different alleles at the androgen-receptor polymorphic trimeric repeat (fig. 5). A total of  $2.6 \times 10^5$  cells were mixed at various ratios and were analyzed according to the method described for the carrier assay. In this case the DNA was not digested with *HpaII* but was treated in the manner described



**Figure 5** Cell-mixing experiment between two males who have different alleles at the trimeric repeat. Cells were mixed in ratios shown to the right of the gels. PCR was performed on undigested DNA from each of these mixtures.

for the undigested control. These results show that, within the range of clinically relevant percentages, the PCR products reflect the starting template concentration.

## Discussion

Differential methylation at the 5' portions of genes on the X chromosome is hypothesized to be one of the primary factors involved in X inactivation (Lyon 1972, 1989; Gartler and Riggs 1983). This methylation has been especially noted at the deoxycytosine residues in CpG dinucleotides (Wolf et al. 1984; Yen et al. 1984; Keith et al. 1986; Hansen et al. 1988; Toniolo et al. 1988). The identification of the methylation-sensitive restriction enzymes *HpaII* and *HhaI*, both of which contain CpG dinucleotides in their recognition sites, has facilitated methylation analysis, since these enzymes only digest the DNA when the deoxycytosine residue is unmethylated. The methylation status of particular genes that contain *HpaII* sites or *HhaI* sites in their 5' regions can be analyzed using these methylation-sensitive enzymes; however, the paternal and maternal copy of the genes must be distinguished from each other in females, if any assessments of methylation are to be made. With the HPRT and PGK genes, RFLPs closely associated with the methylation-sensitive restriction sites in the 5' portion of these genes have been identified, so that the methylation status of both the maternal and paternal X chromosomes can be distinguished (Wolf et al. 1984; Yen et al. 1984; Vogelstein et al. 1987). Another probe in this type of analysis using Southern hybridization in M27 $\beta$ , which detects a VNTR. With this probe, the unmethylated allele is derived from the inactive X chromosome, and the methylated allele is derived from the active X chromosome (Boyd and Fraser 1990; Brown et al. 1990). M27 $\beta$  has not been shown to be associated with any expressed sequences.

The human androgen-receptor locus spans >90 kb of DNA and has been mapped to Xcen-q13 (Kuiper et al. 1989). The cDNA from the gene has an open reading frame of 2,751 nucleotides, from which a protein 917 amino acids in length is translated (Tilley et al. 1989). A unique feature of the gene is the presence of an in-frame CAG trimeric repeat encoding 11–31 glycine residues in the first exon. This repeat is highly polymorphic, with a heterozygosity of 90%, and 20 alleles associated with the repeat have been identified, corresponding to 11–31 CAG repeat units (Edwards et al. 1992). Variations in the length of this repeat

recently have been implicated in X-linked spinal and bulbar muscular atrophy (LaSpada et al. 1991). The fortuitous proximity of two *HpaII* restriction sites (20 bp and 60 bp proximal) and of two *HhaI* restriction sites (37 bp and 78 bp proximal) adjacent to the trimeric repeat led us to test whether the methylation status of the restriction sites correlated with X inactivation.

We report in this paper the correlation between methylation of the *HpaII* and *HhaI* sites in the human androgen-receptor locus and X inactivation. This gene has previously been shown to be affected by X inactivation (Meyer et al. 1975); however, no convenient methylation-sensitive enzyme sites that correlate with the X inactivation status at this locus have been reported. The close association of the enzyme sites with the repeat enabled the methylation status of the enzyme sites on both the maternal and paternal X chromosomes to be determined at a very high frequency in females, because of the highly informative nature of this repeat when a convenient and rapid PCR assay was used. Samples tested include DNA from hybrids containing an active or inactive human X chromosome, females displaying both random and nonrandom patterns of X inactivation, and normal males. We found that both *HpaII* sites and both *HhaI* sites were consistently methylated on the inactive X chromosome, while one or both of the *HpaII* sites and one or both of the *HhaI* sites were unmethylated on the active X chromosome. Each enzyme site was not analyzed separately because the primers were designed to flank the repeat and all four enzyme sites simultaneously. The presence of two, rather than only one, of each of the methylation restriction-enzyme sites is advantageous, since there could be some slight variation, in methylation patterns, between X chromosomes. Having two targets—rather than one—for the methylation involved in X inactivation may insure a more consistent result.

This assay may be applied in cases where nonrandom X inactivation is suspected in individual females. Situations in which this could arise include clonal neoplasms, skewing of inactivation in particular tissues of females who are carriers for various X-linked diseases, carrier females manifesting an X-linked recessive disease, and carrier females who do not manifest an X-linked dominant disease. Although we have shown that the methylation of the *HpaII* and *HhaI* sites correlates with X inactivation, using both enzymes in the assay does not increase the information gained, compared with using just one of the enzymes.

The utility of the androgen-receptor assay relies on the high heterozygosity associated with the polymorphic CAG repeat and on the ease and sensitivity of the PCR reaction. These two characteristics make the assay complementary to previously published assays that detect X inactivation patterns. The combined heterozygosity of the PGK probe (30% heterozygous), HPRT probe (18% heterozygous), and M27 $\beta$  probe (>90% heterozygous) is  $\sim 94\%$ ; the addition of the 90% heterozygosity associated with the androgen-receptor locus pushes the combined heterozygosity of the four assays to >99%. However, this combined heterozygosity is only valid in those situations that provide enough DNA to perform both Southern analysis and PCR. If tissue availability is limited to amounts that can only be studied using PCR, then the recently published PCR assay using the PGK locus (Singer-Sam et al. 1990; Gilliland et al. 1991) along with the androgen-receptor assay gives a 93% combined heterozygosity. Without the androgen-receptor assay, the heterozygosity available through PCR assays would be that of the PGK locus (30%).

The semiquantitative nature of both the Southern and PCR assays may lead to some complications in interpreting X inactivation patterns that are not perfectly random or 100% skewed. This is especially apparent with the two conflicting samples noted above. In each of the conflicting situations, X inactivation patterns are neither perfectly random (50:50) nor perfectly nonrandom (100:0). Whether this is due to actually conflicting results with each of the assays or to quantitative limitations due to the nature of each of the assays is unknown. There is also the possibility of allele-specific competition, where one allele may be more easily amplified than another. We have not conclusively observed this phenomenon, but it may play a role in future experiments. Figure 5 addresses the quantitative potential of the assay and shows the quantitative possibilities available with this assay. Studies designed to address the quantitation problem are ongoing and are leading to a more precisely quantitative and automated approach in determining exact percentages of skewing in X inactivation patterns. This is being approached by using a fluorescently labeled primer in the amplification and then analyzing the products on a device used in automated sequencing.

The sensitivity and semiquantitative nature of the PCR reaction are especially useful in the development of a carrier diagnosis for XLA. XLA, first described by Bruton (1952), is a disorder in which the B lympho-

cytes of affected individuals are unable to differentiate or develop (Rosen et al. 1984). The defect has been found to be cell autonomous, and this is dramatically illustrated in carriers who show 100% skewing of X inactivation in their B lymphocytes: the entire population of B lymphocytes in carriers contains the XLA defect on the inactive X chromosome (Conley et al. 1986; Fearon et al. 1987*b*; Conley and Puck 1988). The difficulty in obtaining sufficient quantities of purified B lymphocytes has made X inactivation analysis using Southern analysis impractical. Carrier analysis assessed through fusion of B lymphocytes from potential carriers with HPRT-negative hamster cells in order to assess X inactivation patterns (Conley et al. 1986; Fearon et al. 1987*b*; Conley and Puck 1988) has been shown to yield consistent results; however, the labor and cost have restricted application of the assay. The carrier analysis reported here has been shown to be fast, easy to interpret, and consistent with results obtained by linkage analysis. However, the separation procedure for the B lymphocytes must insure purity so that contaminating cells will not give false results.

A potential difficulty in interpretation of all X inactivation assays is the occurrence of skewed inactivation via stochastic or nonselective mechanisms. In the case of XLA, the T lymphocytes provide a reasonable control, since they arise from a common marrow precursor and have roughly similar population dynamics. The finding of skewed X inactivation in the T cell lineage toward the same allele in a potential XLA carrier would indicate that the B cells were not selectively involved. The analysis would therefore be uninformative for that individual. Vogelstein et al. (1985) found 3/81 normal females with significant skewing in peripheral blood leukocytes by Southern analysis. Adequate controls are therefore required for each individual studied. We are currently collecting data on skewing in B cells in a large control population to determine whether this frequency varies in the more restricted cell lineage.

## Acknowledgments

Technical assistance by Carolyn Fill is greatly appreciated. We are also grateful for critical discussions with Dr. Al Edwards. This study was supported by the Howard Hughes Medical Institute and National Institutes of Health grants R29 HD22880 (to J.W.B.) and HD24234 (to H.Y.Z.). J.W.B. is an Assistant Investigator for the Howard Hughes Medical Institute.

## References

- Allen RC, Belmont JW (1992) Dinucleotide repeat polymorphism at the DXS178 locus. *Hum Mol Genet* 1:216
- Boyd Y, Fraser NJ (1990) Methylation patterns at the hypervariable X-chromosome locus DXS255 (M27 $\beta$ ): correlation with X-inactivation status. *Genomics* 7:182–187
- Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF (1991a) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349:38–44
- Brown RM, Fraser NJ, Brown GK (1990) Differential methylation of the hypervariable locus DXS255 on active and inactive X chromosomes correlates with the expression of a human X-linked gene. *Genomics* 7:215–221
- Brown CJ, Lafreniere RG, Powers VE, Sebastio G, Ballabio A, Pettigrew AL, Ledbetter DH, et al (1991b) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* 349:82–84
- Brown CJ, Willard HF (1990) Localization of a gene that escapes inactivation in the X chromosome proximal short arm: implications for X inactivation. *Am J Hum Genet* 46:273–279
- Bruton OC (1952) Agammaglobulinemia. *Pediatrics* 9:722–728
- Conley ME, Brown P, Pickard AR, Buckley RH, Miller DS, Raskind WH, Singer JW, et al (1986) Expression of the gene defect in X-linked agammaglobulinemia. *N Engl J Med* 315:564–567
- Conley ME, Puck JM (1988) Carrier detection in typical and atypical X-linked agammaglobulinemia. *J Pediatr* 112:688–694
- Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12:241–253
- Fearon ER, Hamilton SR, Vogelstein B (1987a) Clonal analysis of human colorectal tumors. *Science* 238:193–197
- Fearon ER, Winkelstein JA, Civin CI, Pardoll DM, Vogelstein B (1987b) Carrier detection in X-linked agammaglobulinemia by analysis of X-chromosome inactivation. *N Engl J Med* 316:427–431
- Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 132:6–13
- Fraser NJ, Boyd Y, Brownlee GG, Craig IW (1987) Multiallelic RFLP for M27 $\beta$ , an anonymous single copy genomic clone at Xp11.3-Xcen [HGM9 provisional no. DXS255]. *Nucleic Acids Res* 15:9616
- Gartler SM, Riggs AD (1983) Mammalian X-chromosome inactivation. *Annu Rev Genet* 17:155–190
- Gilliland DG, Blanchard KL, Levy J, Perrin S, Bunn HF (1991) Clonality in myeloproliferative disorders: analysis by means of the polymerase chain reaction. *Proc Natl Acad Sci USA* 88:6848–6852
- Gomez MR, Engel AG, Dewald G, Peterson HA (1977) Failure of inactivation of Duchenne dystrophy X-chromosome in one of female identical twins. *Neurology* 27:537–541
- Goodship J, Lau YL, Malcolm S, Penbrey ME, Levinsky RJ (1988) Use of X chromosome inactivation analysis to establish carrier status for X-linked severe combined immunodeficiency. *Lancet* 1:729–733
- Grant SG, Chapman VM (1988) Mechanisms of X-chromosome regulation. *Annu Rev Genet* 22:199–233
- Greer WL, Kwong PC, Peacocke M, Ip P, Rubin LA, Sini-Novitch KA (1989) X-chromosome inactivation in the Wiskott-Aldrich syndrome: a marker for detection of the carrier state and identification of cell lineages expressing the gene defect. *Genomics* 4:60–67
- Hansen RS, Ellis NA, Gartler SM (1988) Demethylation of specific sites in the 5' region of the inactive X-linked human phosphoglycerate kinase gene correlates with the appearance of nuclease sensitivity and gene expression. *Mol Cell Biol* 8:4692–4699
- Keith DH, Singer-Sam J, Riggs AD (1986) Active X chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5' end of the gene for phosphoglycerate kinase. *Mol Cell Biol* 6:4122–4125
- Kuiper GG, Faber PW, Van Rooij HC, van-der-Korput JA, Ris-Stalpers C, Klaasen P, Trapman J, et al (1989) Structural organization of the human androgen receptor gene. *J Mol Endocrinol* 2:R1–R4
- Kwan S-P, Kunkel L, Bruns G, Wedgwood RJ, Latt S, Rosen, FS (1986) Mapping of the X-linked agammaglobulinemia locus by use of restriction fragment-length polymorphism. *J Clin Invest* 77:649–655
- Kwan S-P, Terwilliger J, Parmley R, Raghu G, Sandkuyl LA, Ott J, Ochs H, et al (1990) Identification of a closely linked DNA marker, DXS178, to further refine the X-linked agammaglobulinemia locus. *Genomics* 6:238–242
- LaSpada AR, Wilson EM, Lubahn DB, Harding AE, Fishbeck KH (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular dystrophy. *Nature* 352:77–79
- Lyon MF (1972) X-chromosome inactivation and developmental patterns in mammals. *Biol Rev* 47:1–35
- (1989) X-chromosome inactivation as a system of gene dosage compensation to regulate gene expression. *Prog Nucleic Acid Res Mol Biol* 36:119–130
- Malcolm S, de Saint Basile G, Arveiler B, Lau YL, Szabo P, Fischer A, Griscelli C, et al (1987) Close linkage of random DNA fragments from Xq21.3-22 to X-linked agammaglobulinemia. *Hum Genet* 77:172–174
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Meyer WJ, Migeon BR, Migeon CJ (1975) Locus on human X chromosome for dihydrotestosterone receptor androgen insensitivity. *Proc Natl Acad Sci USA* 72:1469–1472

- Migeon BR, Moser HW, Moser AB, Axelman J, Sillence P, Norum RA (1981) Adrenoleukodystrophy: evidence for X linkage, inactivation, and selection favoring the mutant allele in heterozygous cells. *Proc Natl Acad Sci USA* 78: 5066–5070
- Moseley AB, Huston DP (1991) Mechanism of *staphylococcus aureus* exotoxin inhibition of immunoglobulin production. *J Immunol* 146:826–832
- Puck JM, Nussbaum RL, Conley ME (1987) Carrier detection in X-linked severe combined immunodeficiency based on patterns of X chromosome inactivation. *J Clin Invest* 79:1395–1400
- Rosen FS, Cooper MD, Wedgwood RJP (1984) The primary immunodeficiencies: first of two parts. *N Engl J Med* 311: 235–242
- Singer-Sam J, LeBon JM, Tanguay RL, Riggs AD (1990) A quantitative HpaII-PCR assay to measure methylation of DNA from a small number of cells. *Nucleic Acids Res* 18: 687
- Tilley WD, Marcelli M, Wilson JD, McPhaul MJ (1989) Characterization and expression of a cDNA encoding the human androgen receptor. *Proc Natl Acad Sci USA* 86: 327–331
- Toniolo D, Martini G, Migeon BR, Dono R (1988) Expression of the G6PD locus on the human X chromosome is associated with demethylation of three CpG islands within 100 kb of DNA. *EMBO J* 7:401–406
- Vogelstein B, Fearon ER, Hamilton SR, Feinberg AP (1985) Use of restriction fragment length polymorphisms to determine the clonal origin of human tumors. *Science* 227: 642–645
- Vogelstein B, Fearon ER, Hamilton SR, Preisinger AC, Willard HF, Michelson AM, Riggs AD, et al (1987) Clonal analysis using recombinant DNA probes from the X-chromosome. *Cancer Res* 47:4806–4813
- Wolf SF, Jolly DJ, Lunnen KD, Friedman T, Migeon BR (1984) Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. *Proc Natl Acad Sci USA* 81:2806–2810
- Yen PH, Patel P, Chinault AC, Mohandas T, Shapiro LJ (1984) Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes. *Proc Natl Acad Sci USA* 81:1759–1763